

A Bias-Free Assessment of Hedge Fund Industry Size, Performance, and Flows

Daniel Barth

Federal Reserve
Board of
Governors

Juha Joenvaara

Aalto University

Mikko Kaupilla

Aalto University

Russ Wermers

University of
Maryland

*Views and opinions expressed are those of the authors and do not necessarily represent official positions or policy of the Federal Reserve Board of Governors, Federal Reserve System, OFR or Treasury.



Aalto University
School of Business

Introduction

Biases in Hedge Fund Data

- ▶ Hedge funds are important participants in many financial markets:
 - ▶ Promote price efficiency through arbitrage
 - ▶ Provide liquidity to illiquid markets
 - ▶ Provide investors unique investment opportunities not spanned by traditional factors
- ▶ The systemic importance of hedge funds was established by the failure of Long Term Capital Management in 1998
- ▶ Recently, many have speculated that hedge funds were integral to disruptions in Treasury markets in March 2020
- ▶ Yet, empirical analyses of the industry are fraught with challenges

Biases in Hedge Fund Data

- ▶ Virtually all publicly reported hedge fund data are voluntarily reported, which leads to various biases in available data:
 - ▶ Selection bias, backfill bias, and survivorship bias are common examples
- ▶ These biases confound estimates of fundamental industry statistics such as aggregate size, performance, and flows
- ▶ For instance, **selection bias** significantly complicates estimates of total industry assets under management (AUM):
 - ▶ Eurekahedge **\$2.33 trillion** (Q3 2017)
 - ▶ Hedge Fund Research (HFR) **\$3.21 trillion** (Q4 2017)
 - ▶ eVestment **\$3.25 trillion** (Q4 2017)
 - ▶ Barclay Hedge **\$3.54 trillion** (Q4 2017)
 - ▶ Preqin **\$3.55 trillion** (Q4 2017)
 - ▶ SEC Private Fund Statistics **\$3.89 trillion** (Q4 2017)

Biases in Hedge Fund Data

- ▶ These biases also affect measures of industry performance and investor flows:

“The dramatic underperformance of hedge funds is pretty amazing considering the survivorship and backfill biases in the index data that skew hedge fund returns upwards by 3% to 5% per year.”

– Peter Lazaroff, Enterprising Investor blog at CFAInstitute.org, February 24, 2016

- ▶ If performance and growth estimates are biased, so will be estimates of the **flow-performance** relationship

The Hedge Fund Industry: An Incomplete Picture

- ▶ Since 2010, 99 papers in the JF, RFS, and JFE have “hedge fund” in their title or abstract
- ▶ Nearly all of these papers use data from hedge fund data vendors, such as HFR or Lipper TASS
- ▶ But vendor data is voluntarily reported, which leads to various (well-documented) biases:
 - ▶ Selection bias, backfill bias, and survivorship bias are common examples
- ▶ These biases confound estimates of fundamental industry statistics such as its **aggregate size, performance, and flows**

This Paper

- ▶ In this paper, we combine vendor data and U.S. regulatory data to provide bias-free estimates of:
 1. Total industry size
 2. Performance, including: alphas, betas, and persistence
 3. Investor flows and industry growth
 4. The flow-performance relationship
- ▶ Our data constitute the most comprehensive view of the hedge fund industry to date
- ▶ Our data also allow important comparisons between **vendor-listed** and **non-listed** funds
 - ▶ Determines the size and sign of cumulative bias in vendor data statistics

Findings

- ▶ We estimate **2016** global hedge fund net assets exceed **\$5.2 trillion**; gross assets exceed **\$8.5 trillion**
 - ▶ At least 40% larger than next largest estimate
 - ▶ Likely still constitutes an underestimate of true industry size
- ▶ **Negative selection bias** dominates positive biases in performance:
 - ▶ Non-publicly reporting funds earn monthly alphas that are **0.28 – 0.45% *larger*** than publicly reporting funds
 - ▶ Non-public performance is also much more persistent
- ▶ Investor flows are much larger for publicly reporting funds
 - ▶ The true flow-performance relationship is much flatter than what is estimated from public data only

A?

Data

Vendor Data

- ▶ Vendor data are derived from a manual consolidation of seven common providers:
 1. Lipper TASS
 2. Hedge Fund Research (HFR)
 3. BarclayHedge
 4. EurekaHedge
 5. Morningstar
 6. Preqin
 7. eVestment

- ▶ Vendor data are collected through 2016. Includes returns, net assets, average and maximum leverage, fund domicile, and fund strategy, among others

Regulatory (Non-Listed) Data: Form PF

- ▶ Dodd-Frank mandated enhanced regulatory reporting for private funds — primarily implemented in Form PF
- ▶ Advisers with more than \$150 million in regulatory private fund assets provide detailed data annually on the hedge funds they advise.
- ▶ *Large Hedge Fund Advisers* (\$1.5 billion in hedge fund assets) provide additional data on a quarterly basis for each of their *Qualifying Hedge Funds* (\$500M in assets)
 - ▶ While reported annually or quarterly, some fields such as returns and asset class exposures, are reported at a monthly frequency
- ▶ Form PF data includes gross and net assets, gross and net returns, asset class exposures, types of borrowing, counterparty exposures and creditors, and much more

Combined Data

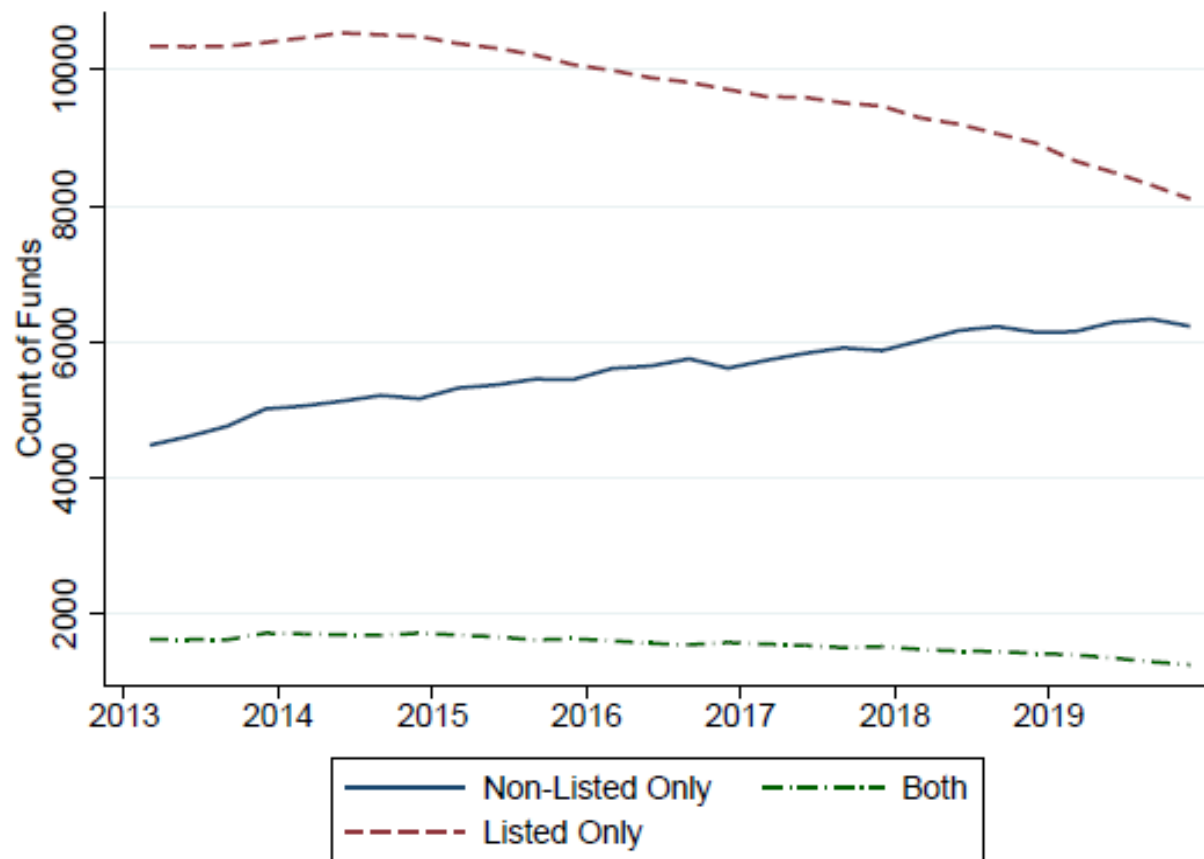
- ▶ To estimate the size of the industry, we must combine data without double counting
- ▶ But, Form PF is **highly confidential**, and vendor data can only be shared with license holders
- ▶ **Our approach:** use Form ADV — a publicly available SEC filing — to get SEC fund identifiers for *each vendor-listed fund*
 - ▶ Then use these identifiers to exclude vendor-listed funds from the Form PF data
 - ▶ This allows us to combine aggregate statistics from the listed and non-listed data without sharing and without double-counting

Resolved and Remaining Biases

- ▶ The Form PF data allow us to observe details about funds that do not report publicly. Funds must report on Form PF if they have:
 1. At least one U.S. investor (excludes family offices)
 2. Advisers with at least \$150 million in private fund assets
- ▶ Form PF resolves biases that result from voluntary reporting: survivorship, backfill, incubation, etc.
- ▶ Our public data also capture:
 1. UCITS (but not 40 act alternative mutual funds)
 2. All but two Billion Dollar Hedge Fund Firms (Edelman et al 2012)
- ▶ While we capture the vast majority of hedge fund assets, we may still miss:
 - ▶ Small funds that don't report publicly
 - ▶ Non U.S. funds that don't report publicly

A?

Combined Data: Fund Counts



How Large is the Hedge Fund Industry?

- ▶ EurekaHedge **\$2.23 trillion**
- ▶ Barclay Hedge **\$2.37 trillion**
- ▶ Hedge Fund Research (HFR) **\$3.02 trillion**
- ▶ eVestment **\$3.00 trillion**
- ▶ Preqin **\$3.14 trillion**
- ▶ *SEC Private Fund Statistics **\$3.48 trillion**

Figure 2: Hedge Fund Industry Net Assets

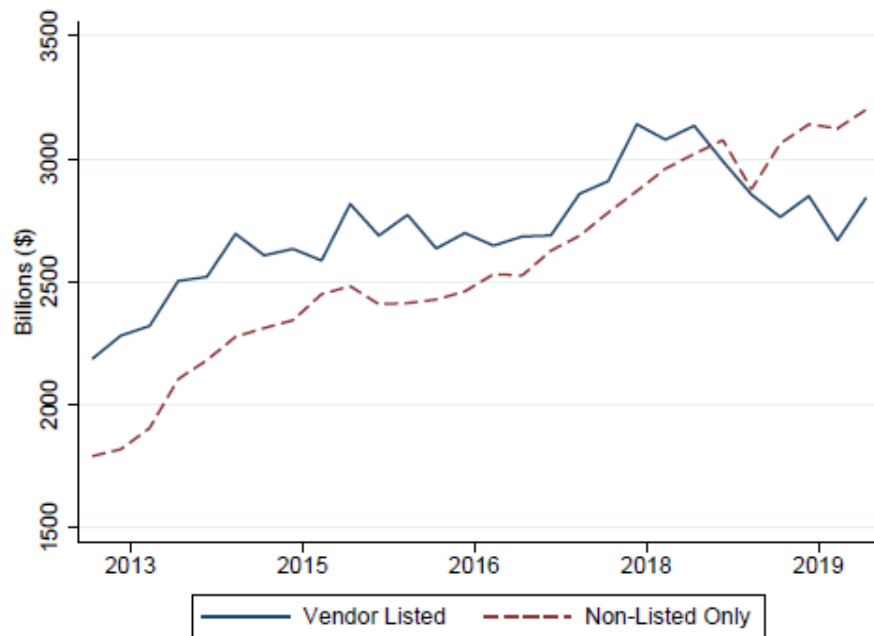


Figure 4: Hedge Fund Industry Gross Assets

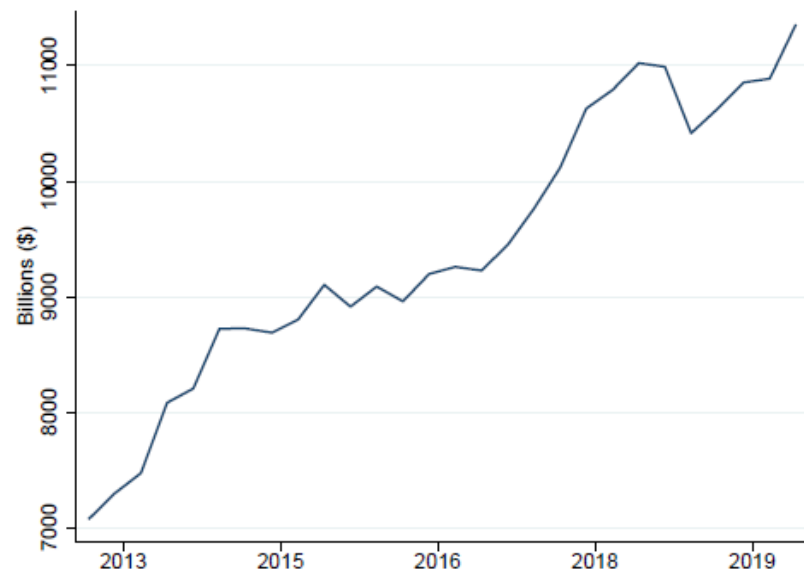
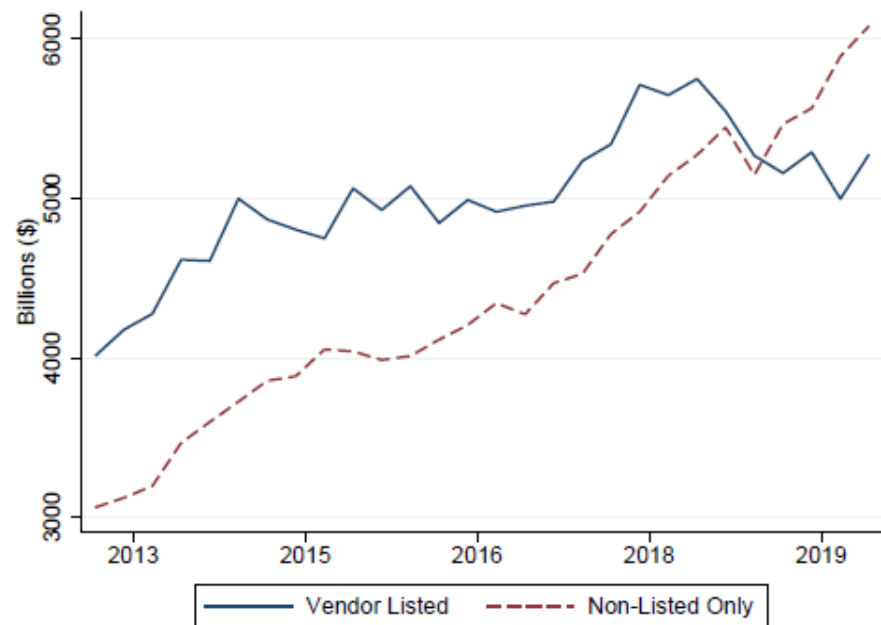


Table 2: Net Assets Under Management by Strategy (\$ Billions)

Strategy	2017			2018			2019		
	Non- Listed Only	Listed	Total	Non- Listed Only	Listed	Total	Non- Listed Only	Listed	Total
Equity	1,342	1,736	3,078	1,324	1,620	2,944	1,618	1,662	3,280
Other	1,108	136	1,244	1,184	135	1,319	1,338	161	1,499
Multi-strategy	1,257	622	1,879	1,162	552	1,714	1,171	450	1,621
Relative Value	236	701	937	284	638	922	320	628	947
Macro	277	1,108	1,384	250	954	1,204	266	957	1,222
Credit	240	762	1,002	425	798	1,223	736	863	1,599
Event Driven	397	292	689	462	274	736	571	274	845
Managed Futures	60	357	417	57	298	355	62	283	346
Total	4,917	5,713	10,630	5,148	5,268	10,416	6,082	5,277	11,359

Table 4: Net Assets by Geographic Region (\$ Billions)

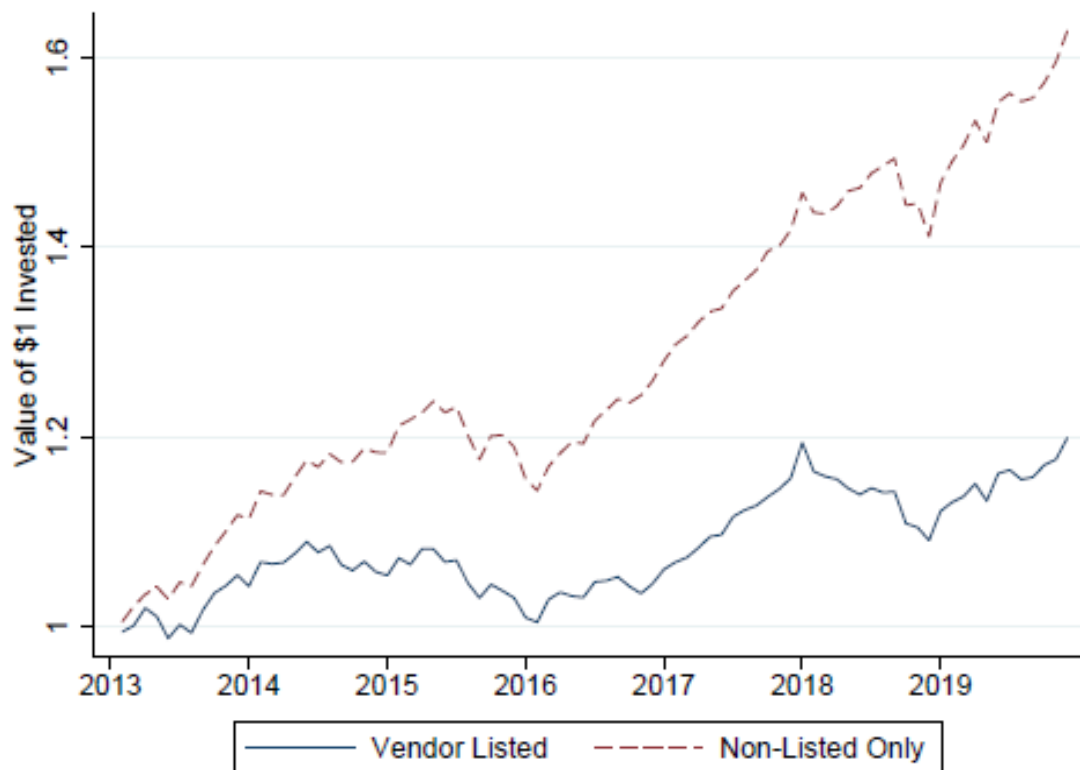
Region	2017			2018			2019		
	Non-Listed Only	Listed	Total	Non-Listed Only	Listed	Total	Non-Listed Only	Listed	Total
Caribbean	1,348	1,016	2,364	1,345	910	2,256	1,523	892	2,415
North America	1,117	855	1,972	1,115	858	1,973	1,252	877	2,129
Europe	123	1,147	1,270	148	969	1,118	160	961	1,121
Others	281	123	404	268	117	385	266	113	379
Total	2,869	3,141	6,010	2,877	2,855	5,731	3,201	2,843	6,044

Resolved and Remaining Biases

- ▶ Many previous studies of performance biases in vendor data suggest the bias is *positive*
- ▶ Edelman, Fung, and Hsieh (2013) find that vendor data are mostly representative; Aiken et al. (2013) and Agarwal et al. (2013) instead find positive selection bias
- ▶ In contrast, we find selection bias is **large**, **negative**, and **dominates** other positive biases

Performance of Listed and Non-Listed Funds

Figure 6: Cumulative Returns



Performance by Strategy

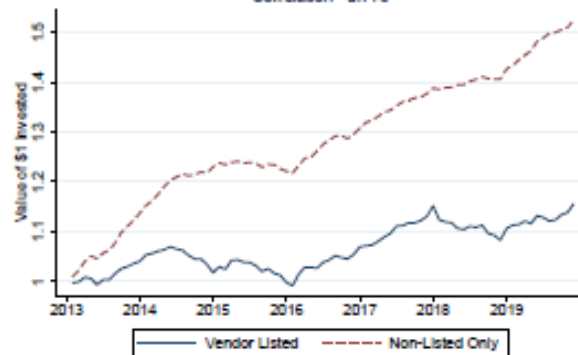
Equity

Vendor Listed: Mean Return -0.411% , S.D. -2.052%
Non-Listed: Mean Return -0.591% , S.D. -1.315%
Correlation -0.913



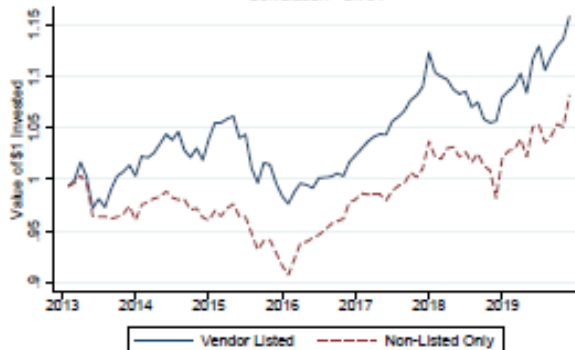
Relative Value

Vendor Listed: Mean Return -0.193% , S.D. -0.892%
Non-Listed: Mean Return -0.591% , S.D. -1.315%
Correlation -0.773



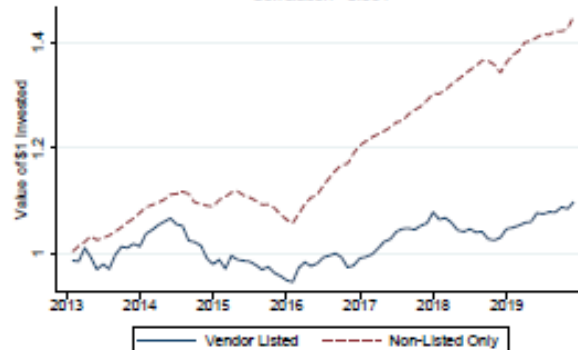
Macro

Vendor Listed: Mean Return -0.195% , S.D. -1.222%
Non-Listed: Mean Return -0.591% , S.D. -1.315%
Correlation -0.791



Credit

Vendor Listed: Mean Return -0.135% , S.D. -1.158%
Non-Listed: Mean Return -0.591% , S.D. -1.315%
Correlation -0.564

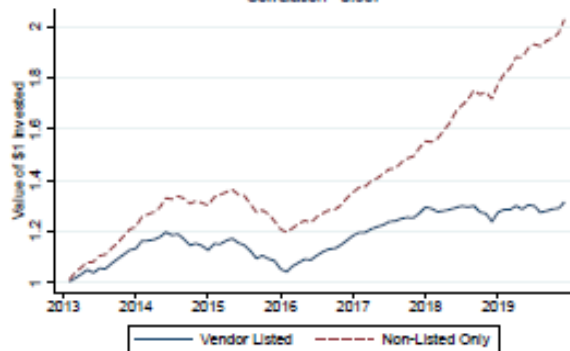


A?

Performance by Strategy

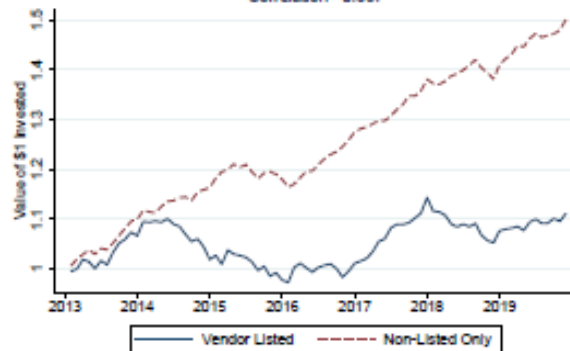
Event Driven

Vendor Listed: Mean Return -0.362% , S.D. $=1.281\%$
Non-Listed: Mean Return -0.591% , S.D. $=1.315\%$
Correlation -0.857



Multi-strategy

Vendor Listed: Mean Return -0.166% , S.D. $=1.303\%$
Non-Listed: Mean Return -0.591% , S.D. $=1.315\%$
Correlation -0.667



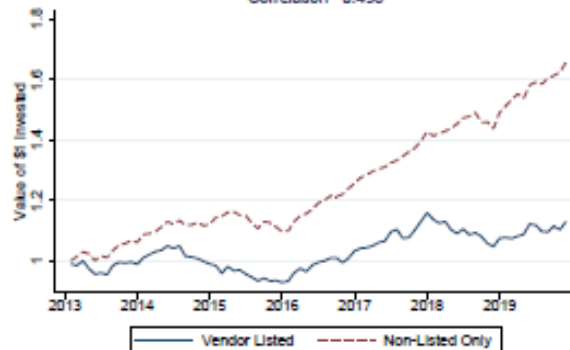
Managed Futures

Vendor Listed: Mean Return -0.129% , S.D. $=1.954\%$
Non-Listed: Mean Return -0.591% , S.D. $=1.315\%$
Correlation -0.262



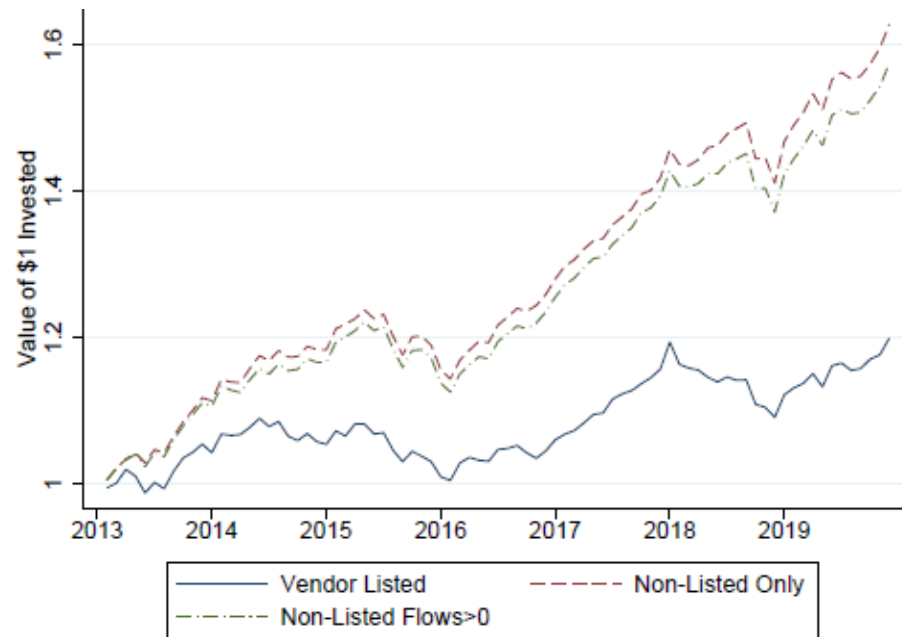
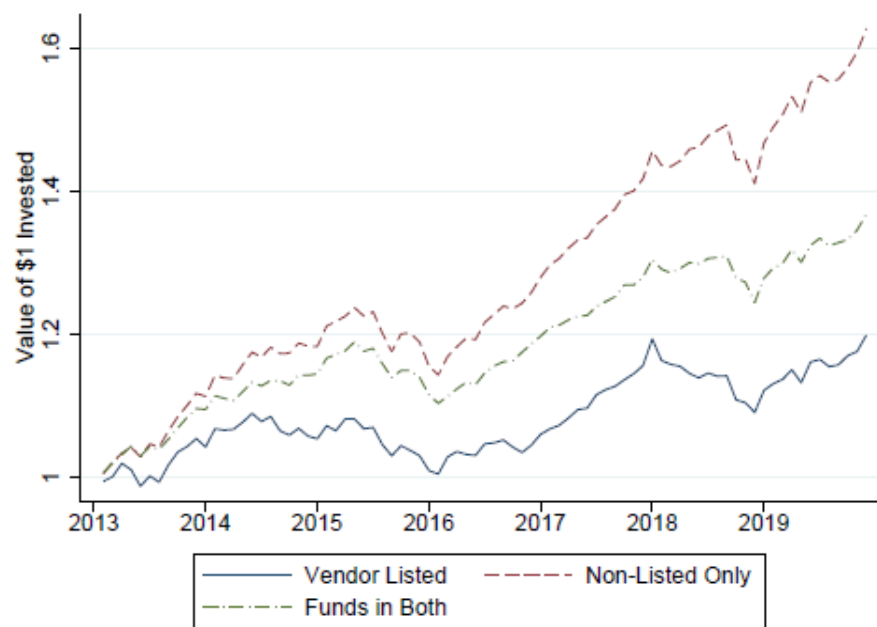
Other

Vendor Listed: Mean Return -0.183% , S.D. $=1.536\%$
Non-Listed: Mean Return -0.591% , S.D. $=1.315\%$
Correlation -0.495



A?

Figure 8: Cumulative Returns: Alternative Explanations



Risk or Risk-Adjusted Returns?

- ▶ Superior performance by non-listed funds could result from two sources:
 1. Greater exposure to systematic risk factors
 2. Greater risk-adjusted performance (alpha)
- ▶ To disentangle these sources, we estimate standard factor regressions of the form:

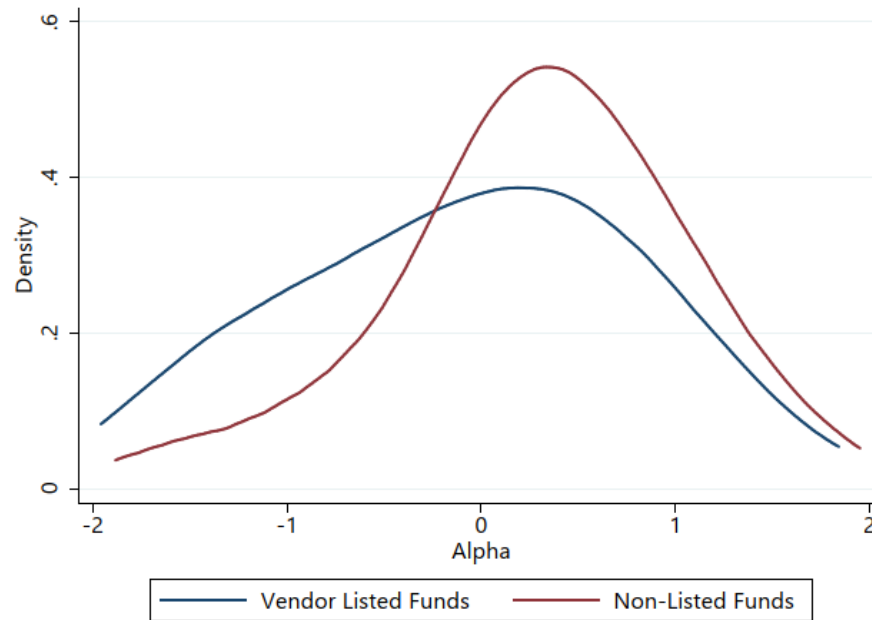
$$R_{i,t} = \alpha_i + \beta_i' F_t + \varepsilon_{i,t},$$

- ▶ We use the “Global 7” factor model from Joenväärä et al. (2019): Mkt, SMB, HML, CSMOM, TSMOM, BAB, PSLIQ
- ▶ We find risk exposures are **highly similar** between listed and non-listed funds [Betas PDF](#)

Measuring Alphas

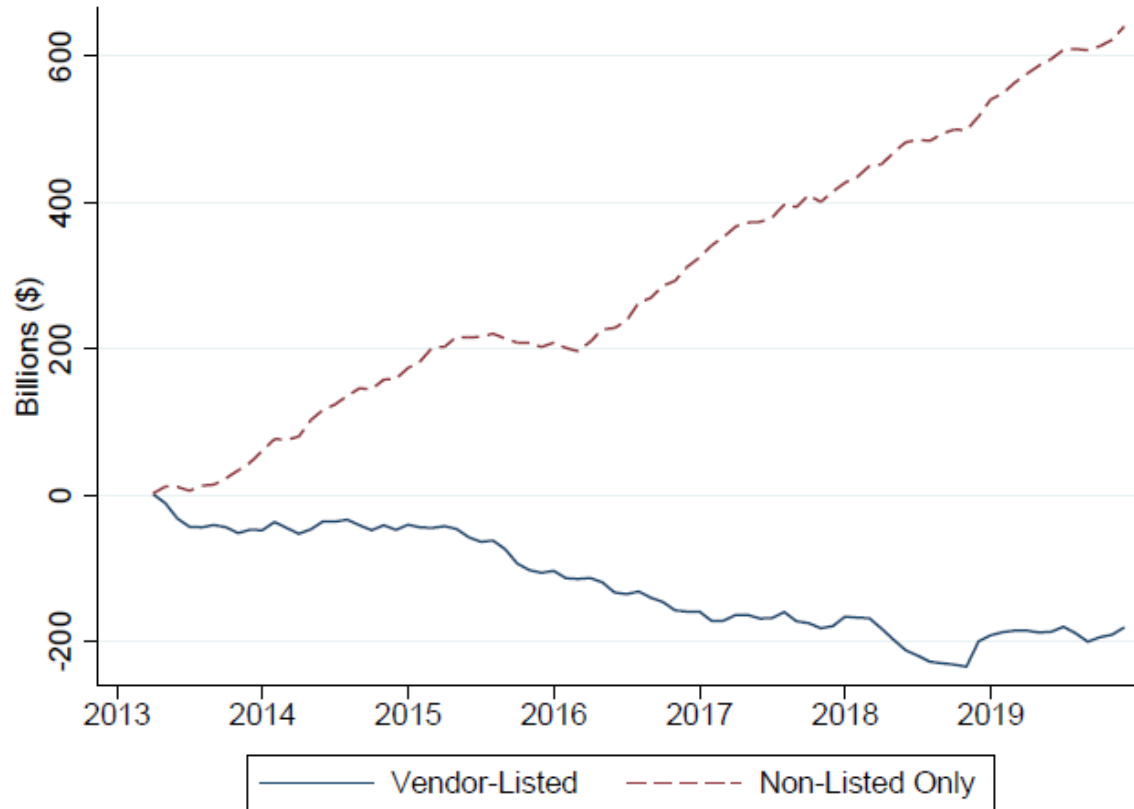
- ▶ Vendor-listed and non-listed funds have similar factor exposures, implying alphas are the source of performance differences
- ▶ We measure the difference between listed and non-listed fund alphas through three approaches
 - ▶ Simple Jensen's alpha comparisons between listed and non-listed hedge funds
 - ▶ Bootstrap (non-parametric inference) of Fama-French (2010)
 - ▶ Fama-MacBeth regressions (control for characteristics of hedge funds)

Jensen's Alpha (Monthly)



- ▶ Vendor-listed funds: mean = -0.146%, median = -0.086%
- ▶ Non-listed funds: mean = 0.470%, median = 0.354%

Figure 11: Berk and van Binsbergen (2015) Value Added



Can Observables Explain Performance Differences?

- We estimate two-stage Fama-MacBeth regressions:

$$\text{(First Stage: TS)} \quad R_{i,t} = \beta_{0,i} + \beta_i' F_t + \varepsilon_{i,t}; \quad \hat{\alpha}_{i,t} = \beta_{0,i} + \varepsilon_{i,t}$$

$$\text{(Second Stage: CS)} \quad \hat{\alpha}_{i,t} = \gamma_{0,i} + \gamma_{1,i} \text{Listed}_{i,t} + \phi' Z_{i,t} + \epsilon_{i,t}$$

$$\text{where } \gamma_1 = \frac{\sum_t \gamma_{1,t}}{T}, \quad (\text{w/ Newey-West s.e.})$$

- We are interested in γ_1 — whether the inclusion of controls attenuates the lower performance of vendor-listed funds

Standard Controls Do Not Explain Differences in Alpha

Dep. Var.	Net Excess Return	Gross G7 Alpha	Net G7 Alpha	Net G7 (GLM Adj.) Alpha	Net FH Alpha	Net FH + Em Mkt Alpha	Net FH + Option Alpha
Listed	-0.38 -4.24	-0.41 -7.20	-0.45 -7.97	-0.45 -6.63	-0.37 -4.31	-0.28 -3.17	-0.34 -3.98

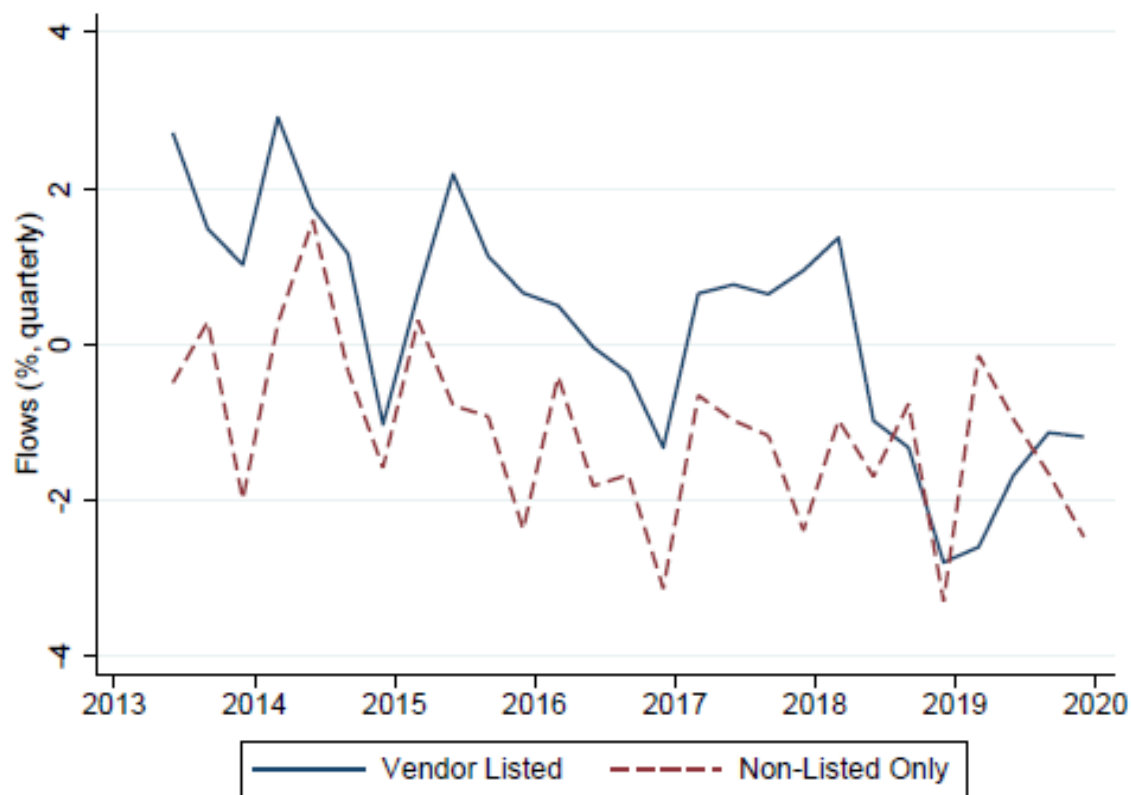
Performance Persistence

Estimation Horizon		3 Months			6 Months			1 Year		
Prediction Horizon		Non-Listed	Listed	$\hat{\gamma}_t$	Non-Listed	Listed	$\hat{\gamma}_t$	Non-Listed	Listed	$\hat{\gamma}_t$
3 Months	Estimate	0.516	0.090	-0.423	0.641	0.176	-0.463	0.677	0.131	-0.545
	<i>t</i> -statistic	10.339	1.971	-6.981	7.603	2.762	-7.707	11.063	0.963	-6.348
6 Months	Estimate	0.507	0.086	-0.420	0.618	0.130	-0.488	0.660	0.099	-0.561
	<i>t</i> -statistic	7.361	2.640	-6.870	6.513	2.026	-9.818	7.984	0.482	-6.188
1 Year	Estimate	0.513	0.063	-0.453	0.611	0.087	-0.528	0.658	0.119	-0.545
	<i>t</i> -statistic	11.688	1.842	-10.382	17.294	0.566	-12.661	21.544	0.457	-5.948
2 Year	Estimate	0.405	0.057	-0.358	0.488	0.107	-0.392	0.537	0.025	-0.520
	<i>t</i> -statistic	6.137	0.906	-12.389	7.666	0.796	-9.497	9.426	0.645	-19.084

► Strong evidence for persistence is found only in non-listed funds

Flows

Figure 13: Aggregate Flows



Flow-Performance Relationship

- ▶ Non-listed reporting funds simultaneously have:
 1. Superior performance
 2. Smaller (proportional) net flows
- ▶ Together, these facts suggest the much-researched flow-performance relationship may be biased upwards
- ▶ To test this, we estimate Fama-MacBeth regressions with flows on LHS, and a vendor-listed dummy on RHS (w/ additional covariates)

Flow-Performance Relationship in Hedge Funds

Dep. Var.	Flow (% Qtr)	Flow (% Qtr)	Flow (% Qtr)	Flow (% Qtr)	Flow (% Qtr)	Flow (% Qtr)
List	4.01	3.89	-0.33	-4.28	-0.09	-0.38
	10.16	8.24	-0.85	-6.38	-0.27	-0.84
List x Performance Rank				0.08		
				12.01		
Performance Rank (percentile)			0.08	0.01		
			17.38	2.87		
List x Net Excess Return						0.27
						5.15
Net Excess Return (% , pq)					0.20	-0.04
					10.11	-0.81
Strategy Controls	No	Yes	Yes	Yes	Yes	Yes
Leverage Controls	No	No	Yes	Yes	Yes	Yes
Offshore Indicator	No	No	Yes	Yes	Yes	Yes
Number of observations	207434	207434	162891	162891	162891	162891
Number of Cross-sections	27	27	26	26	26	26

A?

Economic Interpretation

- ▶ Non-listed reporting funds simultaneously have:
 1. Superior performance, generated through higher alphas
 2. Weaker association between flows and performance
- ▶ Our interpretation:
 - ▶ Skill is scarce, uncertain, and difficult to signal
 - ▶ Managers with uncertain or unproven skill list w/ vendors to generate interest
 - ▶ Managers with established/more certain skill do not need to list
 - ▶ Because of selection, listed funds have zero alpha on average, while non-listed funds have positive alphas
 - ▶ Because listed funds have less certain skill, the association between performance and flows is stronger

Conclusion

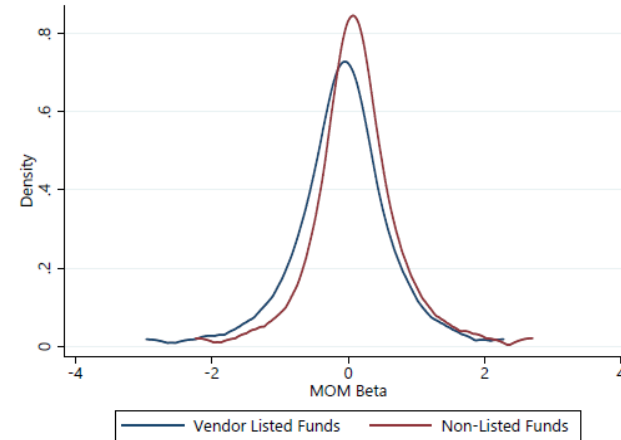
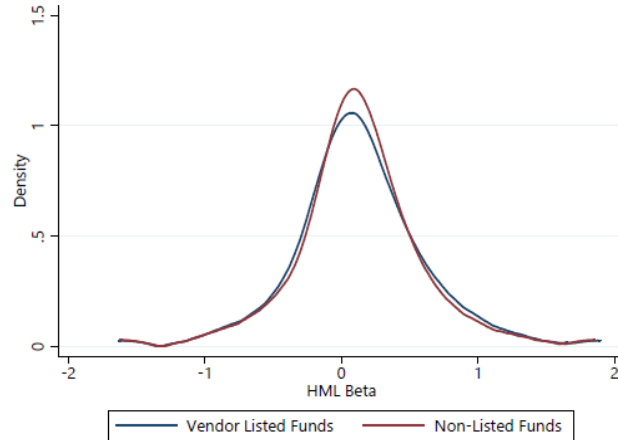
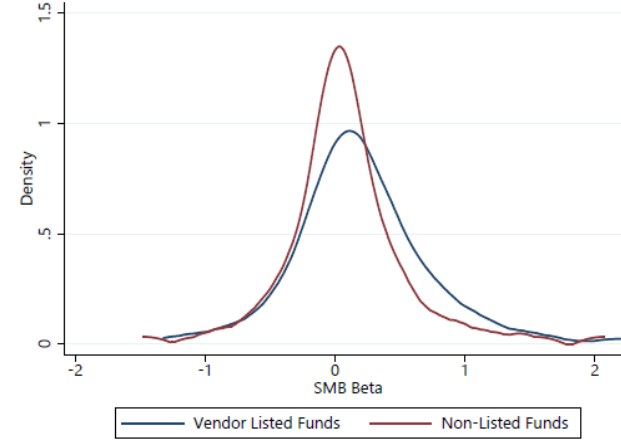
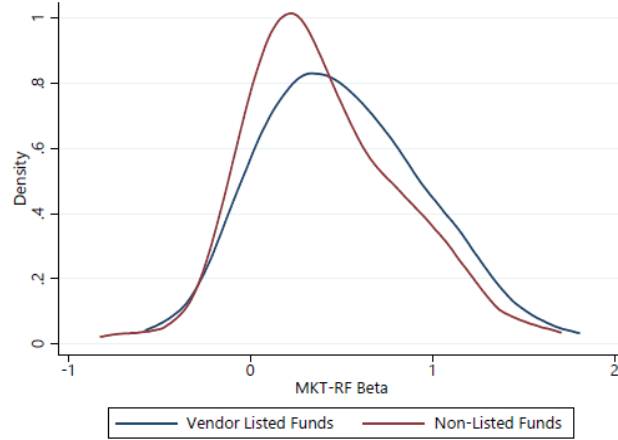
Conclusion

- ▶ We find the HF industry is much bigger than previously thought
- ▶ Vendor-listed funds have much lower performance despite similar exposure to systematic risk
- ▶ Non-listed funds' outperformance comes almost entirely from alpha, and shows significant persistence
- ▶ The flow-performance relationship is much weaker than implied by only vendor data
- ▶ The hedge fund industry is much larger than previously thought
- ▶ The hedge fund industry generates much higher levels of alpha than previously thought

A?

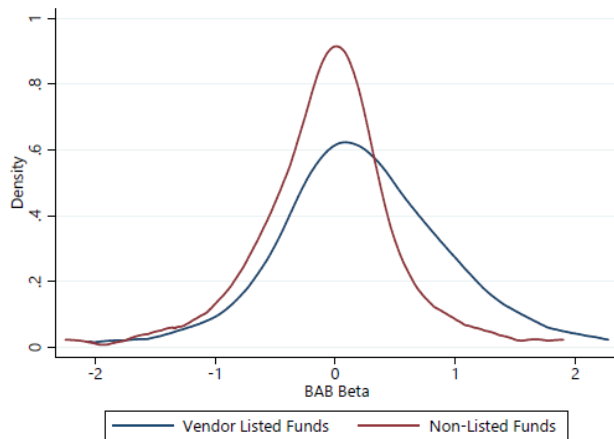
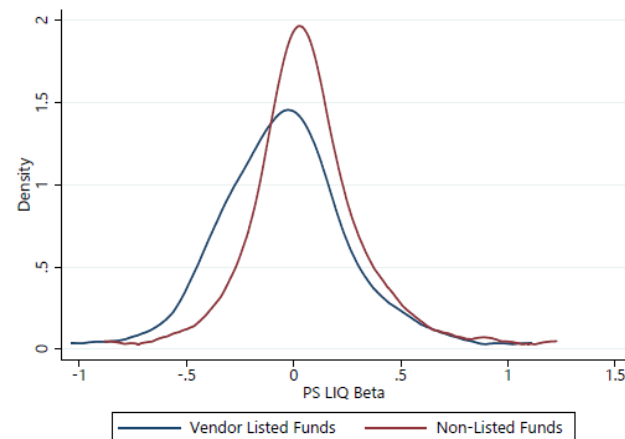
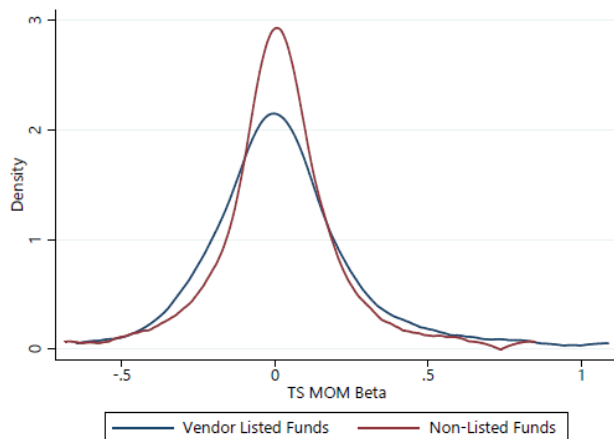
Appendix

Betas on Systematic Risk Factors are Similar



A?

Betas on Systematic Risk Factors are Similar

[Back to Main Slides](#)

A?